R E P O R T R E S U M E S

ED 003 584 56

AN INVESTIGATION OF CONCEPT DEVELOPMENT IN ELEMENTARY SCHOOL SCIENCE TEACHING BY TELEVISION.

BY- GARRY, RALPH AND OTHERS
BOSTON UNIV., MASS.

REPORT NUMBER NDEA-VIIA-527 PUB DATE 1 DEC 63
GRANT OEG-719115-01
EDRS PRICE MF-$0.09 HC-$1.68 42P.

DESCRIPTORS- GRADE 5, *SCIENCE INSTRUCTION, *TELEVISED INSTRUCTION, TELEVISION TEACHERS, *TELEVISION RESEARCH, *PROBLEM SOLVING, TEACHING TECHNIQUES, NATURAL SCIENCES, PROGRAM EVALUATION, *CONCEPT FORMATION, BOSTON, MASSACHUSETTS

THE OBJECTIVE OF THIS STUDY WAS TO TEST THE EFFECTS OF SPECIFIED CONDITIONS OF TELEVISION PRESENTATION AND OF CLASSROOM ACTIVITY UPON THE ACQUISITION OF MEANINGFUL CONCEPTS AND GENERALIZATIONS OF CERTAIN NATURAL SCIENCE PHENOMENA. THE SUBJECTS CONSISTED OF 36 FIFTH-GRADE CLASSES (APPROXIMATELY 1,000 CHILDREN) SELECTED FROM 90 CLASSES WHICH HAD PARTICIPATED IN A PREVIOUS STUDY OF THE TEACHING OF NATURAL SCIENCES BY TELEVISION. TWO TYPES OF TELEVISION PROGRAMS WERE PRODUCED--INFORMATION-GIVING AND PROBLEM-SOLVING. TEACHERS FOR THE EXPERIMENTAL GROUPS HAD BEEN TRAINED IN THE USES OF THESE APPROACHES. THE TEACHERS IN A CONTROL GROUP RECEIVED NO SPECIAL TRAINING. FOLLOWING A TELEVISION SERIES OF 20 PROGRAMS EMPLOYING THE INFORMATION-GIVING APPROACH AND A SERIES (IDENTICAL SUBJECT CONTENT) ORIENTED TOWARD PROBLEM-SOLVING. TESTS WERE ADMINISTERED TO GAUGE THE EFFECTIVENESS OF ONE METHOD OVER THE OTHER. DIFFERENCES IN OUTCOMES OF THE EXPERIMENTAL AND CONTROL GROUP WERE NOT SIGNIFICANTLY DIFFERENT. THE PROBLEM-SOLVING FORMAT RESULTED IN A HIGHER DEGREE OF INTERACTION BETWEEN THE TELEVISION INSTRUCTOR AND STUDENT. OTHER OBSERVATIONS IN FAVOR OF THE PROBLEM-SOLVING PROCEDURE ARE DISCUSSED. (JC)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

AN INVESTIGATION OF CONCEPT DEVELOPMENT IN ELEMENTARY SCHOOL SCIENCE TEACHING BY TELEVISION

Authors: Homer Dietmeier
A. Cornelia Sheehan
Martin Decker

This research was conducted pursuant to a grant from the United States Office of Education, Department of Health, Education and Welfare, under provisions of Title VII of the National Defense Education Act of 1958.

Grant Number: 719115.01

Principal Investigator: Ralph Garry

December 1, 1963

BOSTON UNIVERSITY
BOSTON, MASSACHUSETTS

ED 003584

TABLE OF CONTENTS

I. INTRODUCTION

Problem 1
Objectives 3
Scope of the study 4

II. REVIEW OF RESEARCH

The nature and development of concepts 5
Use of cognitive processes in learning 6

III. PROCEDURE

Population 8
Teacher assignments 9
Teacher training 9
Television materials 9
Experimental design 12
Testing program 13
Construction of tests 13
Test administration 15
Production of experimental television programs 16

IV. DATA GATHERED AND RESULTS

Pretest results 20
Posttest results 22
I.Q. levels and direction of differences 25
Teacher attitude and achievement gains 29
Science reasoning 31
Administrative facilities and achievement gains 33
Retention results 34

V. SUMMARY AND CONCLUSIONS 36

I. INTRODUCTION

Problem

The research problem is to determine the differential effects on the learning of the natural sciences of two types of television presentation which are correlated to two modes of classroom presentation.

At a time when educational television is burgeoning rapidly throughout the nation, when many states are developing plans for instituting or enlarging present educational television facilities, it seems essential that all possible steps be taken to investigate methods of overcoming television's deficiencies as a teaching medium and of taking fullest advantage of its adequacies.

One of the primary goals of education is that of providing a commonly understood body of concepts which are necessary for the individual to function in his society and for a society to function as a unit. Educational television, because of its capabilities as a mass medium, should therefore be concerned primarily with the transfer of a common body of conceptual information.

Educational psychologists and educators are becoming more and more aware that the assumption that concepts can be acquired through memorization is inappropriate. There is mounting evidence that concepts are not static generalizations, but dynamic phenomena, undergoing constant revision, clarification and extension through experience in their use. Therefore, television should provide the students with opportunities to use concepts and to experience their meaning rather than simply reiterate materials which have been previously learned.

Concepts and generalizations which show the relationships between concepts can be taught in several ways. First, the learning can consist of the memorization of a statement of a principle; this ability to verbalize the characteristics of a concept, however, is not necessarily indicative of understanding or the ability to apply the principle to other instances of the same phenomenon. Second, the acquisition of a concept can take the form of understanding without the ability to verbalize the nature of the concept; understanding in this case must be inferred from the behavior of the learner in relation to the concept. Third, a systematic series of experiences may be provided for the learner who attempts to develop from his observations a generalization which will explain or predict the relationship between stimuli which he has perceived; in this case the cognitive process is an inductive one. Fourth, the learner can be provided with a definition or the statement of a principle or generalization and then be exposed to examples which must be identified as fitting or not fitting the definition; this is the use of the

process of deduction to understand the meaning of a principle.

Educational television whether used for direct instruction or enrichment, has most often concerned itself with information giving. Hence, ready-made concepts are assumed to be assimilated or somehow remembered by the learner without any reference to any existing structure. These statements of facts and ready-made concepts are, unfortunately, all susceptible to undesirable consequences: They may not be accompanied by an understanding of any functional relationships; they are often misinterpreted and misused; they provide little potential for association and transfer; and they are easily and quickly forgotten, since they are particularly susceptible to the inhibitory effects of subsequently learned materials.

One of the major concerns about using television as an educational medium is that it is a one-way channel for communication in that the process of feedback and interaction between the teacher and the student is eliminated. The question of how to compensate for the absence of this essential condition for the learning of concepts (i.e., searching for verification and solving problems actively along with the teacher) is important if the full educational potential of television is to be realized.

One method of alleviating the loss of feedback and interaction has been to institute two-way communications systems by providing an opportunity for the learners to question the television teacher. These systems, however, have shown no differential results in the level of learning.

At a seminar held by the National Educational Association, a group of educators, psychologists, communications specialists and educational television producers discussed the importance of interaction in the learning situation and the implications that interaction had for television. They postulated that rote and motor learning, which they classified as linear-serial learning, required the least inter-action, but that cognitive learning (concept development) which they classified as field learning required the greatest degree of interaction. From this basis they postulated that there were two kinds of interactions which could be achieved by television, which though not having the true reciprocal effects of live interaction could approach it. These were termed quasi-interaction and simulated-interaction. The first of these would be what occurs when the viewer identifies with a person in a film or program and experiences the events being portrayed; and the second would be what occurs on television when an experienced teacher, anticipating the response of the viewer, acts as if the response had occurred, when in reality he has no specific information of the event.

Of these two methods of approximating the teacher-student classroom interaction, simulated-interaction would seem to be the most fruitful approach. This procedure would be more directly under control of the teacher and reactions of students would be more uniform. In order to achieve this condition careful pre-testing of materials and instructional procedures is required in order to maximize probability of given responses by the learners.

Objectives

The objectives of this study were to test several hypotheses concerning the effects of specified conditions of television presentation and of classroom activity upon the acquisition of meaningful concepts and generalizations of certain natural science phenomena by fifth grade children, and the corollary effects upon achievement, interests, attitudes, science reasoning and retention of material. Specifically, an attempt was made to answer the following questions:

1. Do differences in television presentation through (a) the provision of relevant direct experience with the phenomena or (b) the provision of information and ready-made generalizations yield differences in pupil gains?

2. Is it possible to increase the amount of cognitive learning by the use of special techniques in preparing and producing educational television programs?

3. Can concept development be achieved through televised presentation alone?

4. What is the effect upon concept development, information gains and interests, of different types of related classroom activity coupled with different types of television presentation?

5. Are differences in achievment related to pupil age, sex, intelligence, and interests?

6. Do differences in television treatment alone have an effect upon the level of retention of material?

7. Do differences in classroom procedure alone have an influence upon the level of retention of material?

8. Do differences in television presentation in combination with classroom techniques have an effect upon level of retention?

Scope of the study

This study was conducted under the provisions of Title VII of the National Defense Education Act of 1958 (PL85-864). The research included approximately 1000 children in 36 fifth grades selected from a group of 90 randomly selected fifth grades within a fifty mile radius of Boston, Massachusetts which had participated in a prior study. The study required the training of 12 teachers in the use of "problem-solving" instructional materials and the training of 12 other teachers in the use of the "informational" instructional materials. Twelve additional teachers, used as controls, were given "no training" and were told to teach as usual. A testing program, including an intelligence test, an interest test, science reasoning tests, a science information test and a science concept test, was administered in September, 1961, May and September, 1962, in conjunction with the reception of two series of 20 television programs by 36 classes during the period of October, 1961 to March of 1962.

II. REVIEW OF RESEARCH

The nature and development of concepts

Curti (1950), surveying the research on concept development, defines concept as "a present item of experience functioning in a symbolic capacity as a response to any one of a class of objects or relationships having a common characteristic." Her failure to clarify the meaning of "item" gives unfortunate latitude to those who would correlate words with concepts. At another point in her review, Curti proposes a number of characteristics that a concept must have, among them: (1) a concept is built up of repeated experiences, (2) concept development is a gradual process, and (3) logical and flexible concepts cannot be acquired through mere word learning." Leeper reporting on cognitive processes, uses "concept" and "cognition" interchangeably, although implicit in his usage is the conclusion that cognition is the application of concepts to problem solving. Furthermore, Leeper (1951: 748) lists six factors which interfere with the formation of new concepts:

"1. A separation of time between cause and effect (as between mosquito bites and malaria).
2. Perceptual incongruities between the cause and effect (as the difference between flies and maggots, in the old theory of 'spontaneous generation').
3. Already existing habits that cause a misleading organization of the items in the situation (as the knowledge that led early workers on beri-beri to search persistently for the transmission of bacteria from one affected person to another).
4. Relative intangibility of either the cause or the effect, or both (as the elusiveness of wind resistance as a factor in the inefficiency of trains at high speed).
5. The existence of interdependent causes (as the fact that neither sufficient calcium nor sufficient vitamin D can prevent rickets by itself, but that both must be present).
6. Motivational factors that work against acceptance of a conclusion (as when Semnelweiss's fellow obstetricians in the 1840's refused to accept his demonstration that puerperal or child bed fever was due to the transmission of some material from one woman to another by the obstetricians themselves)."

One of the most complete analyses of concept formation and perhaps the most meaningful is that of Vinacke (1952). Acknowledging the complex nature of the process under study, Vinacke points out how the study of concepts is further complicated by the complex interrelationships which exist between concept formation and the nature and development of language,

perception and social interaction, plus by the necessity to work primarily by inference since so little is knownabout the internal processes of the nervous system in this respect. With consideration for the complexity of the problem, Vinacke suggests the parameters of research in this area are as follows:

1. We must clarify what we mean by the process of concept formation.
2. We must differentiate between the process of concept formation and the concepts themselves.
3. We must differentiate the genetic aspects of concept formation from the problem solving aspects of concepts.
4. We must differentiate between concept formation as it is found in children and as it is found in adults.

Another outstanding contribution to the understanding of concepts has been made by Brownell (1950). His approach is founded on the importance of meaning and he proposes that all learning can be placed somewhere on a continuum of meaningfulness, depending on the individual learner. Since, Brownell maintains, concepts are dependent upon meaning, concepts should also lie somewhere on this same continuum. The value of Brownell's idea can be further extended if one separates "completeness" of a concept from "meaningfulness" of a concept, and constructs a dual continuum to handle these interrelated though separate attributes of concepts. This dual continuum permits us to perceive how a child can be taught a relatively complete concept and yet have learned a concept that has little or no meaning.

Use of Cognitive Processes in Learning

Effects of using inductive or deductive cognitive processes in learning have been inadequately investigated, even apart from research on the mass media. From the few attempts to study the effects of different forms of acquisition (e.g. Kirsch, 1958; Corman, 1957) the following hypotheses seem to be warranted:

The understanding of the nature of the concept plus the ability to verbalize its characteristics and/or relationships to other concepts is a more useful form of learning than either of these acquisitions taken singly.

Children taught under conditions which encourage a deductive cognitive process (beginning with a given generalization and identifying examples of the generalization through logical reasoning) perform best when tested on the kinds of materials they used in learning by the deductive process.

Children taught under conditions which encourage an inductive cognitive process (the inference of generalizations from direct experience) demonstrate greater gains in using the generalizations thus derived to interpret new materials or solve other kinds of problems.

Independent discovery of generalizations results in greater transfer to new situations; greater motivation probably due to greater involvement and interaction with the phenomenon or task; interest in the task or problem as such apart from other extrinsic incentives; more comprehensive understanding; an auxiliary effect of "learning how to learn" which results in improved problem solving in general; long-term retention (which, together with transfer potential, decreases as a direct function of the amount of aid given in discovery).

Learning by the inductive cognitive process is likely to require more time, and this method is not equally effective for all students--lower ability students are more likely to profit from a "rule-giving" type of approach.

Films coupled with standard instructional procedure have been shown to produce greater eductive learning with high school students (Rulon, 1933) and greater comprehension with adults (Vernon, 1946).

III. PROCEDURE

Population

The population consisted of 36 fifth grade classes (approximately 1000 children) drawn from the Greater Boston Area. These fifth grade classrooms were selected from 90 classes which had participated in a previous study of the teaching of the natural sciences by television, grant number 719008.09, under the auspices of Title VII of the National Defense Education Act of 1958.

Quoting from the report of that study,[1] the population was selected in the following manner:

> "During June, 1959, superintendents of schools in all towns and cities receiving the Channel 2 signal were invited to participate in the study.....Each superintendent was requested, if interested, to supply the following information: (1) population of the community, (2) number of children in fifth grade, (3) number of fifth grades, identified by building, and (4) the number of fifth grades he would be willing to make available for participation in the research project. Using the criteria of per capita expenditure for education it was determined that the cities were sufficiently similar so that, in sampling, only the ratio of towns to cities (approximately 3:1) need be considered.
>
> Using a table of random numbers, fifth grades by buildings in the cities and towns were selected to make a total of 72 experimental classes (48 from towns, 24 from the cities). An additional 18 classes were randomly selected for controls, using the same sampling procedure, completing the total sample of 90 classes, representing 30 communities.....
>
> About 85 per cent of the original sample made good on the superintendent's commitments. During the group meeting for teachers on September 9... teachers were assigned to the workshops on the basis of their expressed preference, filling the quota of 24 for each workshop, but leaving some vacancies to be filled in the "no special training" and "control" groups. These were subsequently filled by repeating the original sampling procedure."

From this sample of 90 classrooms, 36 classrooms were selected for the current study on the basis of the classroom performance on the science information test. The major criterion was the teacher's apparent effectiveness in influencing pupil achievement. The classrooms for this study were selected from the original sample by taking only those classes which fell within the interquartile range in gains in Science Information in the previous study. Those teachers whose classes were at the extremes in gains in science ahcievement were thus excluded providing a more homogeneous sample of teacher effectiveness for the present study. The purpose of the selection of "average only" teachers was to control the teacher skill variable. Our concern was with the effects of given materials and instructional procedures rather than differences in teacher talent.

The 36 teachers were selected on the same ratio of towns to cities (24 towns, 12 city classrooms) as the original sample.

---

1. Report of research on The Integration of Science Teaching by Television into the Elementary School Program, NDEA Title VII, Grant No. 719008.09 Boston University, Boston, Massachusetts, 1960

Teacher Assignments

Of the total number of 36 teachers, 12 were assigned to problem solving groups, 12 were assigned to information groups, and 12 were assigned to "control" groups. The assignments were made by the experimenters on the basis of observations of teachers during the previous study on the premise that success in using either a problem-solving method or information-giving method is dependent upon personal preferences and style of teaching.

Teacher Training

The teachers who were assigned to the problem-solving and information-giving groups were required to participate in a 3-credit course (C, M, 603E) Natural Science Television Workshop: Teaching Methods. The 12 teachers who were participating in the "Control" group were given no specific training, and were told to prepare and teach as they ordinarily would.

The two experimental groups met for twenty, two-hour sessions from November, 1960 through May, 1961. The information group met on Mondays from 4:15 to 6:15 p.m., and the problem-solving group met on Wednesdays from 4:15 to 6:15 p.m.

During the sessions the teachers thoroughly learned the background information for the content to be covered in the ten experimental programs (Television Programs Number 4, 5, 6, 7, 12, 13, 14, 18, 19, and 20). As part of their training, the classes also viewed the latest audio-visual aids available for the subjects to be covered, and finally they worked in teams to create the lesson plans appropriate to each of the ten experimental programs, which evolved as workbooks for the teachers to use in the classroom. The Workbooks for problem-solving and information-giving groups contained identical content outlines and outlines of important learnings; however, the methods to be used in presenting these learnings were decidedly different. The problem-solving workbook used "open-ended" techniques to provoke questions and the "need to know" in the student, whereas the information-giving workbook used verbal presentation of information and ready-made generalizations by the teacher.

All materials to be used in the classroom study of each of the ten experimental programs were sorted and placed in kits for each participating teacher and class. These kits included everything that was to be used in the classroom follow-up including birds' nests, sea shells, worksheets, graphs, etc.

Television Materials

The experimental television materials consisted of two series of 20 one-half hour natural science television programs entitled Exploring Nature. Of the twenty programs in each series, ten were identical for both series. The content areas and concepts to be learned were identical for both series. Only the method of presentation differed. One series was broadcast on Wednesday morning during school, the other series was broadcast on Friday mornings.

The first of the series, Series "R" consisted of 20 natural science television programs which had been prepared by the Massachusetts Audobon Society for the study conducted in conjunction with grant number 71908.09 under the auspices of Title VII of the National Defense Education Act of 1958. The programs had been prepared to emphasize the presentation of information, vocabulary, and concepts and generalizations. Essentially a "tell-and-show" presentation was used. For the second series, Series "E", ten programs in the Series "R" were revised, while the remaining ten were used in original form. In the revised programs

the TV teacher and the content remained the same, but the program format was changed, placing emphasis upon the presentation of inductive problem-solving, the raising of questions, and the use of quasi-interaction.

The programs were:	Series "R" (Regular) Content Description	Series "E" (Experimental)
1.	Alive or Not Alive. This program discusses the distinction between living and non-living matter. It is illustrated by model systems which mimic the various characteristics of life--such characteristics as the use of energy, growth, reproduction and response to stimuli are treated.	No change
2.	These are Animals. The characteristics of animals is the central theme of this program. Movement as a means of obtaining food and the use of food as a source of energy are discussed.	No change
3.	Plants. Plants are living organisms capable of trapping the sun's energy and storing it for future use by themselves and other organisms. Plants control both the numbers and the distribution of some sbrts of animals. This program is primarily concerned with the role of plants in the living world.	No change
4.	Animal Tools and Weapons. The principal use of animal tools is to secure food; their use as weapons is secondary. Some animals, however, possess structures that are used solely for defense. The variety of animal tools and weapons is enormous, BUT all are of value--each is selected to do its job well.	Revised
5.	Mammals. The various characteristics of mammals and the advantages that these characteristics confer on their possessors is the central theme of this program.	Revised
6.	How Animals Move. Animals move in various ways. The basis of this movement, however, is the same in almost all animals--the contraction of a muscle against a fixed object. This program illustrates the action of muscle against bone and compares relative advantages to vertebrates and invertebrates.	Revised
7.	Animal Homes. Homes serve for the protection of the young, for survival of the adult and for the storage of food. The nature of the home of the animal that lives in it, and the variety of homes, are almost as diversified as the animal inhabitants. This program emphasizes how homes are used for the protection of the young and why they are necessary.	Revised

8.	Seeds. The nature of seeds, the variety of methods by which they are distributed, their role in the life of the plant and the process by which they germinate, are discussed and illustrated by this program.	No change
9.	Trees. Trees are rather specialized plants. Methods by which they grow, obtain raw materials to make food, and by which they survive the winter are shown. The economic importance of wood and forest practices are included.	No change
10.	Water. All life depends upon water. Its role in the weathering of rocks, the origin of life, and the sustaining of life are considered. That all animals and plants still live in a watery environment as an aggregation of cells is made evident.	No change
11.	Fossils. Fossils are records in the rocks. They are found in various sedimentary rocks and from them we can learn about life that was present long ago. One of the greatest generalizations in biology, the theory of evolution, is largely based on a knowledge of the fossil record. The emphasis in this program is the role of fossils in the theory of evolution. Formation of fossils is explored.	No change
12.	Darwin and Dinosaurs. From the study of fossils we have learned how animals and plants have changed with time. Dinosaurs were animals specialized in their adaptations to the conditions in which they lived. A study of the history of dinosaurs provides a good case history for a summary of evolution.	Revised
13.	Reproduction. One characteristic of living material is that it reproduces, and reproduction is defined as a process by which an animal or plant gives rise to another of its kind. This program is concerned with the general features which are common to all methods of reproduction.	Revised
14.	Populations and Predators. All living things tend to increase in number. The female codfish, for example, lays over a million eggs, and yet the ocean is not filled with cod. "Why?" They are controlled by competition for food and by predation. This program explores the relation between a population of animals and a predator.	Revised
15.	Animals Prepare for Winter. Winter is a time of scarcity for animals. Both food and water are hard to come by, and in addition the cold presents its own special problems. The various ways in which animals meet this challenge are the subject of this program.	No change

16.	Plants in Winter. Plants also have problems during the winter months. Because of the difference in their method of growth and livelihood, they are able to use different means to solve the winter problem. This program deals with the adaptations for survival utilized by plants.	No change
17.	Now You See It--Now You Don't. The story of animal camouflage is a fascinating one. Animals hide to escape being eaten. But, the many ways in which they conceal themselves are as various as the animals themselves.	Revised
18.	Food From The Sea. The ocean is a vast and unutilized garden. The grass and grazing animals all have their counterparts in the sea. Man is just beginning to develop efficient ways of harvesting the crop from the ocean.	Revised
19.	Life in a Drop. Borrow a microscope and place a few drops of pond water on a glass slide. Here you will visit a strange new world. Strange forms of many colors can be seen bumping, jerking, and shooting through the water. These microscopic single-celled organisms were the concern of this program.	Revised
20.	Blood--the Internal Ocean. Most animals carry inside them an ocean of water which serves a multitude of functions. In mammals it is circulated through an amazingly complex system of pipes by a relatively simple pump.	Revised

Experimental Design

The following experimental design was effected:

		Television Presentation	
		Information-Giving	Problem-Solving
Classroom Method	Information giving	6 classes, teachers trained in information giving method	6 classes, teachers trained in information giving method
	Problem-Solving	6 classes, teachers trained in problem solving methods	6 classes, teachers trained in problem solving methods.
	Control	6 classes, teachers received special training	6 classes, teachers received no special training

1. Teaching Method

   a. Twelve teachers used a correlated study guide distributed by the experimenters. This study guide emphasized the use of lectures and demonstration in which vocabulary was presented, information was disseminated and ready-made generalizations and concepts were presented.

b. Twelve teachers used a study guide which also was correlated with the television series. This guide, however, emphasized the use of problem-solving techniques in teaching. The students were encouraged to explore, question and analyze phenomena, and the emphasis was upon pupil activity in solving problems.

2. Presentation of Classroom Material

In both groups materials were presented introductory to the television program for fifteen minutes, and then presented again as a follow-up to the television program for fifteen or more minutes.

Testing Program

During the course of the research five different tests were administered. These were: science interest, science reasoning, science information, science concept, and the Otis Quick Scoring Intelligence Test.

1. The interest test, used in the previous television study, was designed to assess levels of interest in physical and biological science.

2. The science reasoning test, also used in the previous television study, was used to assess levels of reasoning ability in the science area.

3. The science information test was designed specifically for this study to measure the level of factual information which the pupils had achieved in science as presented in this study.

4. The science concept test was devised for this study to assess the formation of and the ability to use concepts related specifically to the programs presented in this study.

5. The intelligence test was used to establish intelligence levels of the groups participating in the study.

Construction of Tests

Two tests, the science information test and science concept test, were constructed for use in this study.

1. The purposes of these tests were:
   a. the concept test was devised to measure the students' ability totake discrete information, logically organize it, an apply it to a solution of a problem not previously encountered in the learning process in class.

   b. the science information test was devised to measure the students' ability to recall information or fact.

2. Item Construction

a. The original pool of items for the two tests was the science information test devised for the 1959-1960 television project. All items from this pool were subjected to an upper-quarter/lower-quarter item analysis to establish level of difficulty and discrimination. "Good" items thus selected were then analyzed by the following criteria for use in one of the two tests.

(1) Concept item criteria

Concept item must be constructed to:

(a) require the application of information and knowledge not previously encountered in the programs, and
(b) satisfy at least one of the following criteria:

1. involve a relationship requiring the use of two or more attributes of an object to satisfy the relationship.
2. involve multiple relationships, e.g. "all rights"
3. require a "because" or an explanation of "why"
4. require selection of the best answer among true alternates

(2) Information Item Criteria

These items shall be restricted to factual information and shall require a response which:

(a) is an identification of what is true or correct about something.

(b) is information about or characteristic of something.

b. Construction of new items

Each new item was constructed on the basis of the above criteria. After construction, each item was tested independently by another researcher against the item criteria. After this, each item was included in a test of 25 to 30 items which was then administered to a group of sixth graders who had previously viewed the original television series. There were a total of 7 subtests for the concept questions and seven subtests for the information items.

After the items (205 information, 195 concept) had been administered, they were all item analyzed on an upper-quarter/lower-quarter basis. Then on the basis of discrimination, difficulty level, appropriateness of alternates and coverage of content and concept, the items were selected for the instruments to be administered in the study.

c. Make-up of the tests.

1. The concept test consisted of 70 items. There were three items for each of the ten programs which was identical in both experimental series, and four items (including an item in picture form) for each of the ten programs unique to the Series "E" programs.

2. The information test consisted of 100 items. There were five items which dealt with each of the 20 programs in the Series "R" programs.

3. In each of these tests, the items were arranged in specific order of difficulty, ranging from a difficulty level of .15 to .95. The items were sorted into five groups of difficulty, then the easiest of each group were placed as items one through five on each test. From that point the next higher level of difficulty was selected for items six through ten, and so on until the most difficult of each group was used in the final series of five questions.

Test Administration

All tests were administered in early October of 1961, prior to the introduction of the participating pupils to the television or classroom materials, to establish a level of existing knowledge or performance.

Again, in April of 1962, all tests except the Otis Intelligence test were administered immediately following the conclusion of the television series to measure any gains in achievement or changes in performance.

In October of 1962, the science information test, the science concept test and the science reasoning test, part III, were administered to assess any changes in levels of knowledge and performance as a result of retention.

Production of Experimental Television Programs

During the fall of 1959 and the spring of 1960 Boston University, School of Education conducted, under a NDEA Grant, a large scale experiment involving the utilization of thirty television programs in the teaching of biological science to students in ninety fifth grade classrooms distributed throughout the Greater Boston area. These thirty programs were designed and produced by subject matter specialists who held advanced degrees in Biology. Each of the programs followed the same basic format: state a biological principle and then illustrate the principle with a number of examples. The research staff felt that although this format, which was based on traditional teaching procedures, might be adequate for the teaching of facts it might lead to memorization of the biological principle rather than to full development of the concept involved. The staff felt that a reversal of the format might be more conducive to concept development. The reversed format would start with illustrations coupled with the asking of basic questions; through reasoning developed in a series of logical steps the program would evolve the biological principle. It was felt that this reversed format would bring about a greater student involvement and through this involvement a deeper understanding of the biological concept being studied.

In 1960 the research staff applied for, and received, a grant to determine if this reversal of format would in practice lead to better concept development. The second study was limited to the first twenty of the original thirty programs. The first twenty were chosen in order to avoid conflict with school testing programs which ordinarily start in late April. Ten of the twenty programs were selected for experimental purposes, first because they offered excellent opportunity for application of the principle-illustration format and secondly because the programs were organized into three logical blocks: animals, evolution and importance of water or fluid in preserving life. The titles of the ten programs were as follows:

Animal Tools and Weapons
Animal Movements
Animal Homes
Mammals
Reproduction
Darwin and Dinosaurs
Population and Predators
Life in the Sea
Life in a Drop
Blood the Internal Ocean

These ten programs were redesigned and produced in keeping with the reversal format, illustrations evolving into a principle. The other ten programs were held aside for an internal control.

The Programs in the original series were produced at the rate of one per week with the production taking place approximately two or three weeks ahead of the actual televising. Outlines and a program by program study guide were prepared during the summer preceding the telecast. The final programs were scripted and produced under the time pressures stated above. It was decided to develop the new series under circumstances that would involve less pressure. This decision was based on the fact that the new series was to involve the same television teacher who had taught the original series. Since the original teacher had been involved in the design of programs, the format was one which was natural for him and one in which he was in full agreement. Since the new format involved a reversal of technique, it was felt that it would require more time to develop the new programs.

Before commencing the redesign of the ten programs, four frame per second film records were made from the video tapes. These films were made so that the production staff could see exactly how a particular sequence was handled visually in the original program. Transcripts were made of the sound tracks from the original programs, so that the staff could tell exactly what had been said. Using these film records and transcripts the production staff made an analysis of the original ten programs.

The production analysis involved a conference which included the television teacher who taught on the original series, the assistant producer, the producer, and the artist from the original series and the educational psychologist in charge of producing the new series. The producer of the original series was used primarily as a consultant and was not present at all of the conferences.

From the beginning it was realized that less information could be presented in the half-hour programs using the new format. Consequently, it was necessary to reduce the amount of information presented in the program and to confine the new program to the development of the most important concepts. One of the major objectives of the conferences was to identify the major concepts. These major concepts were identified by the television teacher and the assistant producer.

The basic structure of the new programs was one that started with some kind of problem. In some cases the problem was stated directly, in others the problem was presented by showing some natural phenomena which seemed to demand some kind of explanation. Starting with this problem the television teacher would, through a series of questions, explore the logical relationships involved in an explanation of the problem or the phenomena. These questions led to basic generalizations involving the concept being learned or presented. In some cases the instructor answered the questions, in some he did not. Sometimes the generalizations were stated, sometimes they were left up to the students. In any case it was the major objective to allow enough time after a question for the classroom student to think and in some cases to state the answer. While this procedure has features which seem to involve the principles of programmed learning these television lessons cannot be considered as programmed learning. The total structure is far more ambiguous than would be permissable in a truly programmed sequence. Also, the basic objective of the technique is somewhat different. In most programmed sequences the objective is to have the student learn a particular response. In the experimental television lessons described here the basic objective was to stimulate thinking. In theoretical terms it is the general difference between association theory of learning (programmed materials) and field theory of learning (the experimental television lessons).

The conferences between television teacher, producer, artist and educational psychologist led to the development of rough outlines from which the teacher would work. Also the conference led to the development of any new visual material that was felt necessary for the new approach. Film clips, art work and other visual materials were also selected from the original series. An extensive trying-out of materials preceded the making of the videotape for broadcast.

Working with the outline and the selected visual materials the television teacher presented the lesson live in a selected classroom. In the classroom the students answered the questions and discussed the problems that were presented with the teacher. The students were encouraged to reason out answers for themselves. These sessions in the classroom took about an hour and a half to present. A record of the entire presentation and the discussion was made on

quarter inch magnetic tape. Immediately after the classroom presentation the staff evaluated the presentation and made recommendations for changes which would improve the lesson. Care was taken to make certain that the majority of students could think about the problems that were presented. On the basis of the evaluation, suggestions for change in presentation were made to the television teacher. Also changes were made in any visual materials which seemed to require greater clarity. The presentation was then made to a second classroom and the entire procedure repeated. The process was repeated with a third classroom. After presentation to the third classroom the teacher was restricted to the time limits of the television lesson itself. The teacher asked questions and the students were requested to think about the answers, but time was not allowed for discussion. Thus the fourth presentation was quite similar to an actual television lesson. The lesson was discussed once more by the staff before the preparation of the television script which was to be used in producing the final televized lesson.

In practice the above outlined procedure required some modification and uncovered some basic problems which need to be reported here, not only because of their intrinsic value but also because of the possible implications in interpreting the results of the study.

In order to insure that the TV lessons would reach the diverse classrooms to which they were being presented, the four classrooms used in the developmental stage were chosen so as to give a broad range. One classroom was chosen from each of the following socio-economic areas: upper middle class, middle class, lower-middle-upper lower class, and a semi rural area. It was very quickly discovered that it was almost an impossibility to develop a lesson that would reach these four groups. Further it was discovered that considerable care had to be taken in utilizing information gleaned from one class in the development of the presentation for the next class.

It was also discovered that there was basic conflict between the television teacher and the educational psychologist as to technique. While the television teacher was extremely cooperative he did not fully agree some of the basic techniques and sometimes found it difficult to develop a pattern while discussing the material with the class. At the request of the television teacher the educational psychologist prepared more detailed outlines. The detailed outlines were more similar to scripts than were the first rough outlines. Modifications were made in these scripts between classroom presentations.

Also, primarily because of time, it was found advisable to make the presentation and ask questions without immediate discussion with the classroom. The reactions of the students were watched very carefully by the staff during the presentations and the discussion with the children was held until after the presentation. This modified method seemed to result in changes which were as beneficial as those made by using the original method for classroom presentation.

It was felt that the original method might have been better had the four classrooms been more alike, the television teacher more familiar with and in agreement with the presentation technique, and were the amount of time available greater.

The experience of developing the television lessons for this experiment raises two questions which have implications for the interpretation of this study and for the application of research to a specific situation. (1) Is it possible to develop any television lesson equally suited to a diverse group of classrooms? (2) Is it possible to get a television teacher to follow a format which he does not fully appreciate or a format which may be at odds with the teacher's personality?

This last question has broad implications for all research findings. Research may indicate that a particular approach is superior. However, there is a possibility that this approach may be integrally involved with teacher personality and that results are obtained only when the teaching personality is in keeping with the method. Attempts of supervisors and television producers to force teachers to handle a classroom in a specific way or utilize television in a specific way may in the final analysis be a waste of time.

This question also has implications for research on teaching methods. In this study it was decided to keep the television personality the same for both programs. Thus teacher personality was held as a constant. However, this possibly impaired the experimental series because of the fact that the technique was new to the teacher and not fully compatible with his personality. It must be emphasized that one personality is not better than another personality. The point here is, that the method of the original series resulted from the instructional procedure with which the instructor felt compatible and that the method in the experimental series grew from the judgment and interests of the educational psychologist who produced the new television lessons. In the experimental series the instructor was forced to re-educate himself, to the use of methods different from those he habitually employed and perhaps approved. Yet to use a teacher more familiar with the technique we could have introduced a major variable, a different teacher. The problems encountered in the production of the television programs indicate that it may not be possible to bring this variable under control by using the same teacher and, further, that teacher personality and teacher method may be inseparably interlocked.

IV. DATA GATHERED AND RESULTS[1]

To ascertain the distribution of intelligence, each child was administered an Otis Beta Quick Scoring Intelligence Test. Table 1 lists mean scores for boys by cell and Table 2 lists mean scores for girls. Tables 1 and 2 reveal that those classrooms watching the Information Giving programs had a higher mean score than those watching the Problem Solving programs. The difference between the boys' groups (Information Giving: 36.0 and Problem Solving: 32.7) was found to be significant at the .05 level. It should also be noted that the cell representing the interaction between the Problem Solving programs and the Problem Solving classroom method shows a low mean score for both boys and girls. While these means are lower than those of other cells, the difference did not reach a level of significance. However, this cell does show lower scores on most measures made in the experiment.

Tables 1 and 2 - Otis Beta Raw Score Means for Boys and Girls by Experimental Treatment.

CLASSROOM METHOD	Table 1 (BOYS)	TV Format: Information Giving	TV Format: Problem Solving	Total
	Information Giving	34.9	32.6	33.8
	Problem Solving	38.0	31.4	34.7
	Control	35.0	34.1	34.6
	Total	36.0	32.7	

CLASSROOM METHOD	Table 2 (GIRLS)	TV Format: Information Giving	TV Format: Problem Solving	Total
	Information Giving	35.9	33.5	34.7
	Problem Solving	37.9	32.4	35.2
	Control	35.8	35.8	35.8
	Total	36.5	33.9	

---

[1] During the experiment it was discovered that one class was viewing the wrong program series. This failure to follow the design made it necessary to remove this class from the experiment, which in turn made it necessary to remove one class from each of the other five cells of the design. The researchers could have (1) pulled one class from each of the other cells by random means, or (2) pulled the class from each cell which would cause remaining classes in each cell to be more closely matched in intelligence. The latter procedure was chosen and all data for one class in each cell were removed, leaving data for five classes in each cell.

An analysis was made of Pretest scores on all information tests and concept tests. Tables 3 and 4 report the mean scores for boys and girls on the Information Test (items covering both the experimental and the regular series programs). Tables 5 and 6 report the mean scores for boys and girls on the Concept Test (items covering both the experimental and the regular series programs). The pretest correlations between Otis Beta raw score and these two tests covering both series of programs for both boys and girls ranged from .52 to .60. None of the differences reported on the Information Test and the Concept Test (Tables 3 to 6) reached a level of significance.

TABLES 3 and 4 - Information Pretest Mean Scores for Boys and Girls by Experimental Treatment.

CLASSROOM METHOD	Table 3 (BOYS)	TV Format: Info. Giving	Prob. Solving	Total
	Info. Giving	32.3	30.8	31.5
	Problem Solving	31.9	28.3	30.1
	Control	28.9	30.9	29.9
	Total	31.0	30.0	30.5

CLASSROOM METHOD	Table 4 (GIRLS)	TV Format: Info. Giving	Prob. Solving	Total
	Info. Giving	31.6	28.4	30.0
	Problem Solving	32.7	28.3	30.5
	Control	29.2	31.8	30.5
	Total	31.1	29.5	30.3

TABLES 5 and 6 - Concept Pretest Mean Scores for Boys and Girls by Experimental Treatment.

CLASSROOM METHOD	Table 5 (BOYS)	TV Format: Info. Giving	Prob. Solving	Total
	Info. Giving	34.4	33.1	33.8
	Problem Solving	39.5	31.4	35.4
	Control	32.3	34.1	33.2
	Total	35.4	32.9	34.1

CLASSROOM METHOD	Table 6 (GIRLS)	TV Format: Info. Giving	Prob. Solving	Total
	Info. Giving	30.4	29.1	29.7
	Problem Solving	30.7	28.4	29.5
	Control	28.4	31.3	29.8
	Total	29.8	29.6	29.7

Tables 7 and 8 report the mean gain scores on the Concept Test covering the concepts emphasized in the experimental series of television programs. Classes in cells to the left of the tables watched the programs using the information-giving format while classes in the cells to the right of the table watched the programs using the problem-solving format. Both series of programs dealt with the same basic

concepts. The tables indicate only slight differences by treatment. Analysis of variances indicated that these differences were not significant, most of the F ratios being less than one.

TABLES 7 and 8 - Concept Test Gains from Pretest to Posttest for Boys and Girls by Experimental Treatment.

CLASSROOM METHOD	Table 7 (BOYS)	TV Format: Info. Giving	Prob. Solv.	Total
	Info. Giving	5.2	5.6	5.4
	Prob. Solv.	6.0	4.6	5.3
	Control	5.2	5.7	5.5
	Total	5.4	5.3	

CLASSROOM METHOD	Table 8 (GIRLS)	TV Format: Info. Giving	Prob. Solv.	Total
	Info. Giving	5.9	5.8	5.9
	Prob. Solv.	6.1	4.4	5.3
	Control	5.0	6.9	5.9
	Total	5.7	5.7	

A comparison of Table 1 with Table 7 reveals that while there are significant differences in the boys' Otis Beta scores, when those who watched the information-giving programs (Otis Beta: 36.0) are compared with those who watched the problem-solving programs (Otis Beta: 32.7) there is practically no difference at all between the two groups in their gain scores on the Concept Test, 5.4 and 5.3 respectively. It may also be noted that the girls watching the problem solving programs showed slightly higher mean gains on the Concept Test (Table 8) than did the girls watching the information giving programs, inspite of the advantage shown for the latter group in Table 2 (Otis Beta). An analysis of covariances might indicate that the group watching the problem solving programs actually made greater gains on the Concept Test than did the group watching the information giving programs, however when matched samples of different IQ levels were compared (see following section) the differences were not statistically significant.

The cell containing classrooms which both watched the problem solving programs and used the problem solving method in the classroom (used problem solving workbooks) has the lowest gain score on the Concept Test (Tables 7 and 8). There were variations in teachers and communities that were noted during the experimental period that indicated in advance that this cell would probably score low. In addition this cell had a handicap in intelligence (Tables 1 and 2) for both boys and girls.

Tables 9 and 10 report the mean gain scores on the Information Test. The gain scores of the classrooms watching the Information Giving programs and those watching the Problem

Solving programs correspond to the differences indicated in intelligence. The gain scores throughout follow what might be expected from the experimental design--those classes receiving information did better than those concentrating on problem solving. While the direction of these differences is what would be expected, the analysis of variances indicates that none of these differences reaches an acceptable level of significance.

TABLES 9 and 10 - Information Test Gains from Pretest to Posttest for Boys and Girls by Experimental Treatment.

CLASSROOM METHOD	Table 9 (BOYS)	TV Format: Info. Giving	TV Format: Prob. Solv.	TV Format: Total
	Info. Giving	13.4	14.3	13.8
	Prob. Solv.	13.4	8.4	10.9
	Control	13.6	11.9	12.8
	Total	13.4	11.5	

CLASSROOM METHOD	Table 10 (GIRLS)	TV Format: Info. Giving	TV Format: Prob. Solv.	TV Format: Total
	Info. Giving	15.6	13.7	14.6
	Prob. Solv.	13.1	13.1	13.1
	Control	12.9	12.2	12.6
	Total	13.9	13.0	

Information Test "R" and Concept Test "R" covered the content and concepts included in ten television programs which were not changed (known as the "regular" series) but were viewed by all classes in the experiment. No workbooks were developed to cover these programs. Each class experienced a brief introduction and follow-up discussion led by the teacher in conjunction with each program. These tests then measure the effects of a kind of internal control procedure. Since none of the classes had workbooks for these programs, their scores might be expected to be more similar. Also since the control classes were looking at all programs in the same way and consequently did not approach the programs covered by the "R" tests any differently than they would programs covered by the items for the experimental or "E" series, we would expect that their scores might be higher on the "R" tests than would be the scores of the experimental classrooms.

The results of the "R" tests are reported in Tables 11 through 15. The variations within the cells for the most part support the above expectations. The fact that the information scores on the "R" test show a reversal (see Table 11) lends some support to the conclusion that the information-giving method used in the classroom increased the scores. It also indicates that the information-giving group may have learned to depend on the workbooks. The differences obtained in gains on the "R" tests (Information and Concept) did not reach a level of significance by main experimental treatments.

TABLE 11 - Comparison of Gains from Pretest to Posttest on Information Test "R" and Information Test "E" for Experimental Groups Using Two Different Kinds of Workbook.

	Info Test "E"		Info Test "R"	
	Boys	Girls	Boys	Girls
Info-Giving Workbook	13.8	14.6	6.8	6.8
Prob. Solv. Workbook	10.9	13.1	7.4	7.9

TABLES 12 and 13 - Concept Test "R" (items covering programs in the nonexperimental series) Gains from Pretest to Posttest for Boys and Girls by Experimental Treatment.

CLASSROOM METHOD	Table 12 (BOYS)	TV Format: Info. Giving	Prob. Solv.	Total
	Info. Giving	3.7	4.8	4.3
	Prob. Solv.	4.8	2.8	3.8
	Control	4.0	5.0	4.5
	Total	4.1	4.2	

CLASSROOM METHOD	Table 13 (GIRLS)	TV Format: Info. Giving	Prob. Solv.	Total
	Info. Giving	4.4	2.2	3.3
	Prob. Solv.	4.7	3.6	4.1
	Control	4.4	6.3	5.4
	Total	4.5	4.0	

TABLES 14 and 15 - Information Test "R" (items covering programs in the nonexperimental series) Gains from Pretest to Posttest for Boys and Girls by Experimental Treatment.

CLASSROOM METHOD	Table 14 (BOYS)	TV Format: Info. Giving	Prob. Solv.	Total
	Info. Giving	6.7	6.8	6.8
	Prob. Solv.	8.0	6.7	7.4
	Control	8.7	10.5	9.6
	Total	7.8	8.0	

CLASSROOM METHOD	Table 15 (GIRLS)	TV Format: Info. Giving	Prob. Solv.	Total
	Info. Giving	8.2	5.4	6.8
	Prob. Solv.	8.8	6.9	7.9
	Control	7.8	8.5	8.2
	Total	8.3	7.0	

Interest Test and Attitude Test results showed that differences by experimental treatment fluctuated around zero (plus or minus) for both boys and girls yielding F ratios of less than one.

IQ levels and direction of differences.

One advantage of laboratory experimentation is the degree of control over all variables provided the experimenter. Difficulty ensues with regard to making generalizations. In contrast, field experimentation suffers from the limited aspects of the environment (stimulus conditions) which can be controlled. In the present study we were able to bring under experimental control approximately one of the twenty-five hours spent in school. This was the complete portion of curricular time allocated to science per se, but certainly not to problem-solving, concept development, or for that matter, information about science. A perusal of the reading series textbooks in use in fifth grades reveals considerable science information presented. Hence extraneous influences may or may not be having an effect on the results of a field experiment to the degree to which transfer of training occurs. These extraneous influences can produce accelerating or counteracting effects. The experimenter is faced with the problem of not accepting chance differences as real, yet not overlooking true differences which are small (Type I versus Type II errors).

In this study, setting the customary levels for rejecting the null hypothesis (the .05 and .01 level of statistical significance) indicates that no significant differences were obtained as a result of the experimental variables, or to be more accurate, the few differences which were statistically significant of the many tested showed no systematic association with the independent variables which permitted attributing them to the experimental design.

One is always faced with the question of whether or not the planned experimental conditions were obtained, in this instance not in respect to the follow-up activities because teachers had been trained and materials developed, but with particular respect to the two sets of TV programs. Two sets of observations verified the fact that the new programs at least had "face validity" for adults shown a pair of programs, one old and one new, could describe the differences in procedure and in their response to the two programs. This occurred regardless of order of presentation and with the only prior information being that the difference was in organization of the programs. Secondly, differences in responses of experimental classes were readily observed. Each class which viewed the ten new programs also saw ten of the original set. There were readily observable differences in the response of the classes - raising hands to answer questions posed by the TV teacher, vocal responses given to questions, higher attention level, in other words greater interaction was obtained between program and pupil.

Two further analyses of the achievement test data were made. The first was to check for differential effect associated with

intelligence level. In selecting pupils of "high" and "low" intelligence, a matched sample was drawn which permitted not only a comparison of achievement by intelligence test level, but also a test of main effects with intelligence controlled so as to reduce the original differences which had been observed between cells. Although there was not anticipation that heretofore insignificant differences would now attain statistical significance, we were interested in the extent to which the direction of differences coincided with a priori predictions. For example, one would anticipate that youngsters experiencing the revised programs and/or a problem solving follow-up should achieve higher scores on the concept test than youngsters experiencing a straight information approach, and vice versa.

In matching the subjects it was necessary to eliminate all subjects with an IQ below 95 and above 130. The low group consisted of subjects from 95 to 110. The high group consisted of subjects from 110 to 130. A total of 48 subjects were selected from each cell (after pooling of all classrooms). Twenty-four subjects were boys and twenty-four subjects were girls of which twelve were in each of the IQ groupings indicated above. Subject cards were selected by hand with considerable care being taken to make certain that approximately equal representation of classrooms was obtained. Classroom representation was given preference over exact matching in those cases where exact matching might have caused over weighting in a particular classroom. Table 16 reports the mean IQ scores for the various cells. An analysis of variances indicated that none of the differences reached a level of significance.

TABLE 16 - Means and standard deviations of intelligence by experimental treatment of matched sample of boys and girls (N = 12 per cell)

	Follow Up	Boys: TV Program Info	Boys: TV Program Prob	Girls: TV Program Info	Girls: TV Program Prob
High Intelligence IQ 110-130	Info	116.1 3.7	116.9 4.5	117.5 4.2	116.2 2.9
	Prob	116.9 4.5	119.5 4.9	117.5 4.2	116.8 3.9
	Control	118.5 4.2	117.1 4.1	116.6 4.6	115.2 3.1
Low Intelligence IQ 95-110	Info	104.8 3.4	104.4 3.6	103.2 3.1	104.7 3.7
	Prob	104.3 3.9	103.9 3.4	104.8 3.7	105.0 4.5
	Control	104.6 4.4	103.7 3.9	105.2 4.8	105.5 3.4

The premises behind the research would also predict that those classrooms using the problem solving workbooks would do better on the concept test than would those classrooms using the information giving workbooks. Further it would predict a higher probability of the control classrooms scoring higher than the classrooms using the information giving workbooks. This latter prediction is based on the assumption that there will be teachers within the control classrooms that will use a problem solving approach as their normal procedure. The prediction of superiority in the problem solving classrooms held in only the Boys Hi IQ group. The prediction of superiority in the control group held in all but one of the groups, Girls Hi IQ.

In the case of the Information Test, by hypothesis one would predict that those classrooms using the information giving workbooks would do best, the control next best and the problem solving classrooms the poorest. The better performance by the information giving classrooms held true in all groups except the Hi IQ boys. The control proved to have the higher score in only two of the groups.

In these instances, eight predictions of high score and eight prediction of low score were made, 16 predictions in all. Only 8 predictions were correct. With a one in three chance of correct prediction the probability of predicting eight correct is .08.

The chance of predicting both Hi and Lo scores in the same IQ group is one in six. Eight such predictions were made and 3 of them were correct. The probability of this occurring by chance is .064.

Table 17 reports the mean differences for the various IQ groupings on the Concept Test E and on the Information Test E. The results reported in these tables are as a whole fairly close to what we might expect as a logical prediction based on the hypothesis underlying the design of the experiment. Those students watching the problem solving television programs would be expected to show greater gains on the concept test than those watching the information-giving programs. Table 17 reports this to be true in all four groups (Boys Hi, Boys Lo, Girls Hi and Girls Lo). The difference in both Boys Lo and Girls Lo was found to be significant at the .05 level. The analysis of the Boys Lo difference revealed an F ratio of 4.33 and the analysis of the Girls Lo difference revealed an F ratio of 6.81. An F ratio of about 4.00 being required for significance at .05 level and a ratio of about 7.08 being required for significance at .01 level.

Also the theory would predict that those students watching the information giving programs should do better on the Information Test than those students watching the problem solving programs. This prediction was true for all groups except the Boys Hi IQ. None of the differences found with the Information Test E were great enough to reach a level of significance.

In seven of the eight instances the differences coincide with predictions. The probability of seven out of eight predictions occurring on a chance basis is .017.

TABLE 17 - Gains made by matched IQ groups on Concept Test (A) and Information Test (B) by Experimental treatment. (N=12 per cell.)
* Asterisk indicates correct prediction.

A. CONCEPT TEST E

BOYS (Hi IQ)

	Info	Prob	
Info	5.1	5.7	5.4*
Prob	7.8	10.0	8.9*
Control	5.8	5.8	5.8
	6.2	7.2*	

GIRLS (Hi IQ)

	Info	Prob	
Info	9.5	7.6	8.6
Prob	8.0	8.4	8.2
Control	6.8	9.4	8.1
	8.1	8.5*	

(Lo IQ)

Info	2.8	4.2	3.4
Prob	0.8	4.9	2.9
Control	2.9	5.8	4.38
	2.2	5.0*	

(Lo IQ)

	2.2	4.3	3.3*
	4.7	4.6	4.6
	1.8	9.0	5.4
	2.9	6.0*	

B. INFORMATION TEST E

BOYS (Hi IQ)

	Info	Prob	
Info	12.4	17.6	15.0
Prob	18.1	19.7	18.9
Control	17.8	16.0	16.9
	16.0	17.8	

GIRLS (Hi IQ)

	Info	Prob	
Info	17.6	19.7	18.6*
Prob	15.2	18.2	16.7*
Control	21.3	14.6	18.0
	18.0*	17.5	

(Lo IQ)

Info	11.4	14.2	12.8*
Prob	11.7	6.3	9.0*
Control	10.6	10.2	10.4
	11.2*	10.2	

(Lo IQ)

	13.7	8.6	11.1*
	10.3	11.8	11.0
	9.3	9.4	9.4
	11.1*	9.9	

* Correct Predictions

The chance of making all predictions correctly within a single IQ group is one in twelve. Out of the eight groups predicted three were predicted correctly for all major design variables: Boys Hi IQ on the Concept Test, Boys Lo IQ on the Information Test and Girls Hi IQ on the Information Test. The probability of making this many correct predictions by chance alone is .0014 or about one in a thousand.

The foregoing analysis represents more of a consolation to the experimenters than a significant finding of the study. There is no question in our minds that pupil ability and teacher skill are more significant factors in achievement than programs. Nevertheless, the indications of differential effects by type of program suggests the potential for interaction between learner and program going beyond the typical passive television presentation.

Teacher Attitude and Achievement Gains.

Each teacher was administered the Minnesota Teacher Attitude Inventory. The results of this test are shown in a graphic form in Table 18. The graph reveals a very clear differentiation between the teachers in the problem solving group and the teachers in the information giving group. It also shows a greater spread of scores covering the entire range for the control classroom teachers. There are two possible explanations of this rather clear cut differentiation between the three groups. Teachers selected for the problem solving group and teachers selected for the information giving group were selected on the basis of whether or not the experimenters, who had seen the teachers working in the classroom, thought that the teachers basic style of teaching fit either the problem solving method or the information giving method. This may account for the differentiation. It is also possible that the differentiation is a by-product of the training period, in that the teachers made an attempt to answer the Minnesota Scale according to what they thought was wanted on the basis of the training they had experienced. The more even distribution found in the control room teachers is undoubtedly due to the fact that they were randomly selected for their role plus the fact that they did not receive any special training. The gains in achievement on both the information test and the concept test by the pupils of these respective groups of teachers was not sufficiently different to yield a significant difference associated with teacher score on the Minnesota Inventory.

TABLE 18 - Graphic Distributions of scores on Minnesota Teacher Attitude Inventory.

Score	Control Teachers	Info-Teachers	Prob-Teachers
70	×		× × ×
60			× × ×
50	× ×	×	
40	×	× ×	× ×
30	×		
20		× ×	
10	× ×	× ×	
0	×	×	
-10			×
-20			
-30		×	
-40			×
-50			

A further comparison between scores on the Information Test E was made using all classroom scores grouped in thirds: Ten classrooms taught by teachers receiving the highest score on the Minnesota scale, Ten classrooms taught by the ten teachers receiving the lowest score on the Minnesota scale, and Ten classrooms made up of those in between. The group taught by teachers with the ten high Minnesota scores had a mean gain score of 14.8, the ten middle classrooms had a mean score of 11.4 and the ten low classrooms had a mean score of 10.2. Again an analysis of variance indicated that the difference was not large enough to be significant; however, the F ratio was 3.25 when a ratio of 3.34 was needed for significance at the .05 level. There is some indication here that there could be a relationship between teachers' scores on the Minnesota scale and Information gains in the classroom.

Science Reasoning

Two of three tests of science reasoning which had been used in an earlier study were given to all classes before and after the experimental period. The first of these is most appropriately thought of as an abstract reasoning test. A child was required to note the similarities and differences in sets of abstract figures and identify the common characteristic of each set. None of the gains on this test reached statistical significance, although the girls who had viewed the problem-solving TV programs made consistently greater gains.

The second test involved problems dealing with biological science. Each question involved a problem question, e.g. Do whales nurse their young? Each question was followed by a set of five facts, some of which were relevant to the problem. All sets included the information necessary to solve the problem, so that its solution was not dependent on or especially aided by prior knowledge. In fact, the questions were selected to present novel situations. Subjects were required to indicate which of three solutions they thought correct, and more important, check the facts needed to prove the solution. This test appears to measure deductive reasoning.

Tables 19 and 20 report the gains scores on both tests. It is to be noted that on the second test the differences between the girls watching the problem-solving TV programs and those watching the information-giving programs are statistically significant in all three sub-groups at the .05 level of confidence. The F ratio of 7.53 approaches the .01 level which requires a ratio of 7.82. However when the scores on both tests are combined, the differences for both boys and girls

TABLE 19 - Mean Gains made by boys and girls (by experimental treatments) on Science Reasoning - Abstract.

BOYS

	Info	Prob	
Info	5.5	4.0	4.7
Prob	4.6	3.6	4.1
Control	5.4	6.4	5.9
	5.2	4.7	

GIRLS

	Info	Prob	
Info	5.0	5.9	4.4
Prob	4.6	5.9	5.2
Control	3.9	5.8	4.8
	4.5	5.9	

TABLE 20 - Mean Gains made by boys and girls (by experimental treatment) on Science Reasoning - Deductive.

BOYS

	Info	Prob	
Info	4.1	5.4	4.8
Prob	7.6	5.0	6.3
Control	6.6	6.7	6.6
	6.1	5.7	

GIRLS

	Info	Prob	
Info	4.1	7.1	5.6
Prob	4.6	7.3	6.0
Control	4.6	8.9	6.8
	4.5	7.8	

fall short of statistical significance. When one checks the pre- and post-test mean scores on the deductive reasoning test alone, one finds that both boys and girls have comparable scores.

	Pre-test	Post-test	Different
Boys	46.8	53.0	6.2
Girls	47.2	53.5	6.3

To accept the difference in the gains made by girls watching the two sets of programs, one would have to assume a sex difference in which the information-programs were having a depressing effect upon the girls growth in reasoning and the problem-solving programs an elevating effect. In spite of the significant F ratio, it is more reasonable to assume it to be a one in twenty chance difference than a significant finding.

Administrative facilities

Near the end of the experimental period each teacher was asked to fill out a form which asked forty-three questions about the school system in which she taught and methods which she used in class. The form was used as an additional check on the distribution of factors that might effect the outcome of the experiment. After a survey of the answers it was decided that several needed to be checked against the gain scores in the classrooms. Items in the form that were checked were as follows:

1) Does your school have a student library?

6) Do you have enough science text books so that each child can have a copy to keep throughout the year?

8) Do your children have a period during the week in which they have free choice of reading material?

(A series of questions leading to a very positive audio-visual program within the school system.)

24) Do you require the children to take notes during the television program?

25) Do you test the children with a written test after they watch the programs?

Of these items only one, number 24, proved to be of importance. No specific instructions had been given the teachers regarding note taking during the program. The question was added to the check sheet when it was observed that some teachers were giving their students instructions on note taking. Note taking seemed to be fairly well distributed throughout the various classrooms in the research design; consequently, the teachers were not requested to cease the practice.

One finding of interest was that a statistically significant difference in gains on the concept test (.001 level) was obtained by both boys and girls on that part of it concerned with the experimental (problem-solving programs) in those classes in which note-taking occurred. The difference was independent of the intelligence level of the pupils. The gains made on the information test were not related to note-taking, nor on the half of the concept test concerned with the regular programs.

Retention of learning

In October, six months following the administration of the achievement tests on which the previous discussion has been based, the information, concept, and science reasoning tests were readministered to determine the amount of forgetting. Through the cooperation of participating schools, the children who by now were in the sixth grade were re-tested in a series of three sessions. In those instances where the classes had not been promoted intact, the participating pupils were reassembled for testing.

Comparison of gains and losses was made on the separate parts of the information and concept tests which related to both the revised and original television programs, so that in effect, there were two part scores for each test plus a total test score. Analysis of variance was applied to differences in mean scores of classes between the April and October testing, and again for the groups of boys and girls matched for intelligence.

The differences between the April and October scores were negligible. In other words, there was very little loss from the gains made during the previous year. A mean loss of approximately one point occurred on the information test, with two-thirds of the classes showing slight losses, one-third slight gains. The situation was just the reverse on the concept test.

None of the experimental treatments proved superior when one looks to class means, sex differences, and IQ differences taken in combination. The only point with either the information test scores or the concept test scores at which statistically significant differences were observed was on informational gains on the experimental television programs. Here high intelligence girls who had seen the information type TV programs retained significantly more, and low IQ girls (IQ95-110) who had viewed the experimental TV programs using the problem-solving approach were significantly better (P of .05 in both instances). Although the boy's scores showed a similar trend, the differences were not statistically significant. In evaluating these results our inclination is to assume they may reflect a simpler and perhaps better organization within the experimental programs which made comprehension easier, rather than any inherent superiority of process. Scores on the Science Reasoning tests showed marked gains for all classes, both experimental and control, which were statistically significant in most instances. This held true for both boys and girls and by IQ level. When one looks to experimental treatments, however, the only instances in which statistically significant differences were obtained were two. The information type program combined with information type follow-up was significant superior to control groups; however the exact reversal is noted with problem-solving programs in which control groups viewing these programs surpassed experimental groups. With the high IQ girls, those viewing the problem-solving approach in TV programs showed significantly greater gains.

One must note the not surprising fact that both boys and girls of higher IQ showed greater gains on the reasoning tests than the lower IQ boys and girls in nearly all cells. In view of the intercorrelation between this test and intelligence test scores (.55), the gains in scores over the six month interval appear to us to be a function of maturation with concomitant gains in mental capacity.

A final comment should be made with respect to patterns that appeared with respect to combinations of program and follow-up and retention. There were six combinations of program and classroom follow-up in operation: a

straight problem-solving approach, a straight informational approach, and four combinations of mixtures. In two of the mixed classrooms followup is controlled in that the teachers were provided with materials and training in their use. In the other two instances (the control classes) the follow-up was left to the discretion of the teacher.

In evaluating the retention results, the differences which reached or approached significance appeared to favor some form of problem-solving approach either in TV programs or follow-up. We, therefore, checked the retention in reference to each of the six combinations just described, noting the number of instances in which a combination yielded highest retention, lowest retention, or fell in between. The two best approaches appeared to be a problem-solving follow-up linked with either type of program. In no instance did the straight problem-solving approach come out worse. The three worse approaches included both types of TV programs with unspecified follow-up (control groups) and straight information.

V. SUMMARY AND CONCLUSIONS

With funds provided under the provisions of Title VII of the National Defense Education Act, Boston University School of Education implemented a research project in the period 1960-1962 involving the production and transmission of an experimental series of television programs on the subject of natural science to 36 elementary school classrooms; the pretesting, posttesting and retention testing of the children in these classrooms; the training of 24 of the teachers of these classes in special instructional methods; and the preparation of special instructional materials to be used in the classrooms in conjunction with the programs.

The major research objective was to test hypotheses concerning the effects upon the attainment, application and retention of concepts of a problem-solving format as opposed to an information-giving format for the televised presentations paired with problem-solving methods and materials as opposed to information-giving methods and materials used in the classroom.

One experimental group was composed of 12 teachers (and their classes) who attended a course designed to train them in the use of a problem-solving approach to the study of natural science phenomena and during which they developed instructional materials consistent with a "discovery" approach; another experimental group was composed of 12 teachers (and their classes) who attended a course designed to train them in the effective communication of information, ready-made concepts and generalizations relevant to natural science phenomena during which they developed instructional materials consistent with a "show-and-tell" approach. Half of the classes in each of these groups viewed the television program series with the problem-solving format and the other half of the classes in each of these groups viewed the television program series with the information-giving format. Twelve teachers (control group) received no special training or instructions as to how to conduct their classes. Half of these classes watched the information-giving television series, the other half watched the problem-solving television series.

The information-giving television series consisted of 20 programs which had been prepared and produced for use in a previous study. The problem-solving television series consisted of 10 old and 10 revised programs especially prepared and produced for use in the present study. The television teacher was the same person for both series, and only program content contained in the original series was included in the revised programs. Both series were identical with respect to subject matter content and the concepts and generalizations to be mastered; only the method of presentation (teaching) differed.

Tests administered to the experimental and control classes to assess achievement effects consisted of an intelligence test; a science information test and a science concept test, both of which were constructed for use in this study; a science reasoning test and a science interest test, both of which had been constructed for use in a previous study. Statistically significant differences in achievement gains as a function of main experimental effects were not evident except in isolated instances. Initial differences in mean I. Q. by experimental group were noted, but when matched samples of different

I. Q. levels were compared for gains by experimental treatment, the differences were still not statistically significant. While the achievement differences by treatment did not reach statistical significance, predictions of the direction and order of change by treatment based on a priori assumptions supported these assumptions.

Marked differences in teacher attitude by experimental group were observed, due largely to the selection of teachers for assignment to these groups based on assumptions about their ability to operate more effectively with a problem-solving or an information-giving approach to natural science. Gains in achievement by their classes, however, were not significantly related to these differences in teacher attitude.

A significant but subsidiary finding was that note-taking by students during the televised presentation was associated with gains by both boys and girls in the attainment of concepts dealt with by the problem-solving television series. The gains were made independently of intelligence level. Note-taking was not associated with gains in information.

Retention testing revealed little loss of gains made from pre- to post-testing. Retention was not significantly different by experimental treatment with the exception of the fact that high I. Q. girls who viewed the information-giving television series and low I. Q. girls who viewed the problem-solving television series retained significantly more information. The boys' scores show a similar but not significant trend. Marked gains in reasoning were observed for all classes, but these gains were not consistently associated with experimental assignment, and may more reasonably be attributed to maturation.

Viewed in one light, the outcomes of the study were discouraging, for no experimenter or educational practitioner desires to put the hours of labor in development of materials and execution of research which ends in no yield. Yet these findings are not inconsistent with certain of the reviews of research in this field, which show that the more rigorous the controls imposed, the fewer the number of significant differences obtained. Although this may be a salutary situation as far as school administrators are concerned because it provides them with options as to what instructional procedure to follow based on other considerations such as staff, cost, availability of materials, etc., we could not view it with the same equanimity.

There was little doubt, in our judgment, that the revised programs and the follow-up materials were yielding observable behavioral difference, at least in the responses of the children. There was a much higher degree of interaction between the television instructor and classes viewing the programs with the problem-solving format. Interest and attention levels were higher. One could discount the achievement tests used as insensitive, but one would have to engage in detailed achievement test administration after each program to get the full measure of effects of differences in treatment. Yet our educational objectives are in the residual effects, not the immediate effects.

In terms of improvement of instructional materials, however, there would seem to be wisdom in the approach employed here. In this instance, we attempted to hold teacher skill constant, both within television programs and in the classroom. Socio-economic differences, intelligence, and

instructional materials were similarly controlled, in order that program format and process could be manipulated as the experimental variable. In spite of these efforts, the fraction of the environment which we were able to bring under the control in a field situation was still so small compared to the total life exposure of the child that our skills were not commensurate with the task undertaken, at least not on an initial attempt on this scale. In our next attempt we believe it would be more useful to extend small-scale probes and trial variations designed to create differences, measuring these for short-term effects, and doing this under less rigorous experimental conditions, yet searching for ways to maximize the yield. After such trial and error learning, such a field study might be more profitably undertaken, instead of following the procedure of moving from given hypothetical constructs to the applied situation. In this way we would, so to speak, identify and purify the essential experimental treatments.

Ultimately our goal is to make maximal use of the various instructional procedures available to use and to identify what might be thought of as optimal mixes of materials and procedures. This necessitates an understanding of the dynamics of the medium and how it functions vis-a-vis the learner as much as how to harness the medium to existing educational objectives. There is no doubt that the objectives we establish are partially determined by the procedures and technology available. Our initial efforts to explore some of the dynamics of the medium, not only in relation to informational gains, but more difficult, the attainment of concepts and change of attitudes, have been partially defeated by our regard for experimental rigor as well as the difficulty in modifying conventional approaches to instruction. In retrospect, all of the gains are not measured by the statistical data, and as is often the case with experimental studies, the next steps are more clearly perceived.